#30
J. Douglas
5/2/03

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Gregory HOUSE

Title: THREE DIMENSIONAL STRUCTURE ESTIMATION APPARATUS

Serial. No.: 08/962,315

Filing Date: 31 October 1997

Examiner: S. AN

Art Unit: 2613

APPEAL BRIEF UNDER 37 CFR §1.192(a)

Commissioner for Patents
Washington, D.C. 20231

Sir:

Applicant hereby appeals the Final Rejection issued 19 November 2002.

1. Real Party in Interest

The present application is assigned to NEC Corporation of Tokyo, Japan.

2. Related Appeals and Interferences

There are no known related appeals or interferences.

3. Status of claims

Claims 1 and 4 are canceled.

Claims 2-3 and 5-18 are pending. All of claims 2-3 and 5-18 are rejected over prior art and all rejections are appealed.

4. Status of Amendments

No amendments have been made after the final rejection.

5. Summary of the Invention

The claimed invention involves determining the distance to an object by triangulation using camera images.

The triangulation method, also referred to as stereo imaging, is described in the background section of the application at pages 2-3 with reference to Figure 6. Briefly stated, the conventional triangulation method utilizes two cameras having different known locations along a baseline. When the cameras are pointed at exactly the same point on an object, the angles of the cameras with respect to the baseline and the distance between the cameras may be used to calculate the distance to that point (page 2, first through third full paragraphs). Ideally, the object is exactly centered in both images and therefore must be located at the intersection of the central axes 603, 604 of the cameras' fields of view 605, 606.

Figure 6 also shows that the intersection of the fields of view of the two cameras is relatively narrow (page 7, last paragraph). In one sense it is desirable for the cameras to have wide fields of view so that the area of intersection of the fields of view is wide (page 7, last paragraph). However, in another sense, narrower fields of view are more desirable, since in a narrow field of view each pixel represents a smaller part of the object, and so the images are more detailed and alignment of the cameras to the same point on the object is more accurate (page 8, last paragraph).

The claimed invention improves over the conventional method by enabling the triangulation method to be performed using one camera that has a wide field of view and another camera that has a narrow field of view (page 9, last paragraph), or using cameras that have different resolutions (page 10, first paragraph). This allows for highly accurate distance measurement across a wide field of view (page 11, last paragraph).

An example of an embodiment using a wide visual field camera and a narrow visual field camera is shown in Figures 1 and 2. In this embodiment, the object appears small in the wide visual field image 5 and large in the narrow visual field image 6. If the cameras have the same resolution, then a pixel of

the wide visual field image 5 represents a much larger portion of the object than a pixel of the narrow visual field image 6. The conventional triangulation method does not provide good results if the images have different pixel units (page 15, second full paragraph), therefore the claimed invention provides a conversion unit that processes the images so that their pixel units are equal in the amount of the object that they represent. In one embodiment, this is done by sampling pixels of the image that has the higher resolution to produce a sampled image having approximately the same resolution as the low resolution image (Figure 1; page 15, last paragraph). In another embodiment, multiple sampled images are produced from the higher resolution image, with each one sampling different sets of pixels of the higher resolution image (Figure 3; page 23, first paragraph through page 24, first paragraph). In another embodiment, the pixel values of the lower resolution image are interpolated to produce an interpolated image having approximately the same resolution as the higher resolution image (Figure 4; page 26, second full paragraph through page 27, second full paragraph). In each case the resulting images are processed using triangulation to calculate a distance to the object.

The independent claims recite applications of the basic features of the invention in various configurations, including a configuration using cameras having different resolutions (claims 2, 5, 7 and 9), and a configuration using a high resolution camera and a low resolution camera, and cameras having different resolutions (claims 3, 6, 8 and 10).

Each independent recites the two following features:

> ***converting*** at least one image outputted from the cameras ***such that the pixel units of all images are equal in the amount of object represented thereby***;

and

> ***processing the images using triangulation*** to calculate a distance to the object.

Therefore all rejections may be overturned if it is found the cited combination of references fails to teach either one of these features.

6. Issues on Appeal

The issues on appeal are:

1) Whether the Auty reference teaches converting images having different resolutions such that their pixel units are equal in the amount of object represented, or teaches determining a distance to an object by triangulation using such images?

2) Whether the Subbarao reference teaches converting images having different resolutions such that their pixel units are equal in the amount of object represented, or teaches determining a distance to an object by triangulation using such images?

3) Whether the combined teachings of Auty and Subbarao would motivate one of ordinary skill to perform triangulation using images that have different resolutions by converting the images such that their pixel units are equal in the amount of object represented despite providing no explicit teaching or suggestion of either of these features?

7. Grouping of Claims

All claims stand and fall together.

8. Argument

All claims were rejected under 35 USC §103(a) as being obvious over Auty (U.S. 5,809,161) in view of Subbarao (U.S. 5,193,124). However, the technologies described in Auty and Subbarao are only marginally related to the technology of the present claims, and central features of the claims are absent from both Auty and Subbarao.

Teaching of Auty

Auty discloses a system for making detailed images of vehicle license plates. As shown in Figure 1, the system uses a first camera 6 for imaging vehicles at a distance, and a second camera 8 for making a detailed image of a vehicle when it is within a fixed range. The respective distance ranges of the

cameras are shown in Figure 3, and the width 12 of the field of the first camera 6 is shown in Figure 4. Images from the first camera 6 are used to track vehicles as they approach the second camera 8 in order to estimate the time at which each vehicle will be in the range of the second camera (col. 4, line 53 - col. 5, line 10). This estimate is used to trigger the second camera to take detailed images of the vehicle, from which license plate information may be extracted (col. 5, lines 10-16).

Auty estimates a distance to a vehicle, however this estimation is made using only the images from the first camera. As explained at col. 7, lines 1-39 and col. 20, line 29 - col. 21, line 63, the position of a vehicle in the image plane of an image from the first camera is converted into a position on the X - Y plane of the road surface (see Fig. 4) based on the fixed relationship of the camera relative to the road surface. In other words, based on where the vehicle is in the image, it can be determined where the vehicle is along the road. Auty does not use its two cameras to produce two different images of an object so that the distance to the object can be computed by triangulation as claimed. In the advisory action dated 15 April 2003 it is asserted that Equation (1) at col. 7 of Auty teaches triangulation, however col. 7, lines 7-17 plainly state that Equation (1) calculates the distance along the road that is covered by the first camera 6 based on the height of the camera above the road, the distance to the nearest portion of the road within the camera's field of view, and the lens field of view angle. Further, at col. 20, line 29 - col. 21, line 63, Auty explains how the distance of vehicles being tracked by the single first camera is determined by converting the coordinates of a vehicle's location within the image plane into the real world coordinates of the vehicle along the road surface. Therefore Auty does not teach the determination of the distance to an object by triangulation.

The official action further notes Auty's equation at col. 21, lines 1-10, and asserts that since this equation is stated by Auty to represent a "transform which performs scaling, translation and perspective correction," then it is the same as the claimed conversion by which the pixel units of all images are made to represent the same amount of an object. However, examination of the cited portion of Auty shows that Auty's Equation (14) (col. 21, line 1) is an alternate

form (see col. 13, line 67) of Auty's Equation (13) (col. 20, line 50), which is plainly stated by Auty as being used to convert the coordinates of a vehicle's position within the image plane into the real world coordinates of the vehicle's position along the road surface, as discussed previously. This conversion has nothing to do with making the amount represented by pixels of the images of one camera equal to the amount represented by of pixels of the images of another camera.

Therefore, while Auty's system includes two cameras that have different resolutions, this is the only similarity to the claimed invention. Auty uses a single camera to determine the distance to a vehicle, and uses another camera to make a detailed image of the vehicle. Auty does not use triangulation to determine the distance to a vehicle, and Auty does not convert the images made by one camera so that their pixels represent the same amount of an object as pixels of another camera.

Teaching of Subbarao

Subbarao discloses a system that determines the distance to an object for purposes of better focusing a camera on the object (col. 3, line 65 - col. 4, line 3). Subbarao determines distance by making two or more images of the object along the same line of sight, with each of the images being different with respect to one or more of the distance between a principle imaging plane (i.e. a lens) and an image detector, the camera aperture, the focal length, or spectral characteristics (col. 4, lines 15-28). The two or more images are processed to determine a distance to the object (col. 4, lines 28-32). Subbarao's preferred embodiment uses a single camera with switchable lenses that enable the camera to make images having different focal lengths (Figure 1; col. 14, lines 33 - 41). In an alternative embodiment, Subbarao's system uses two cameras with two different focal lengths (Figure 4; col. 16, lines 5-24). However, those cameras produce images made along the same line of sight, not from different viewing angles. As shown in Figure 4, the light emanating from the object "Scene" is received at a half-silvered mirror, which distributes it to both of cameras 1 and

2. It is clear as shown in Figure 4 that the image produced by each of the cameras is produced from exactly the same viewing angle.

Subbarao therefore does not determine distance through triangulation. Because Subbarao's system uses only one camera or two cameras that image along the same line of sight, Subbarao does not produce images of the object from different viewing angles as claimed. Therefore it is not possible for Subbarao to determine the distance to an object by triangulation, since triangulation requires the use of images made from two different visual angles, knowledge of those angles relative to some reference, and knowledge of the distance between the two locations at which the two images are made, so that computations using those angles and distances can be made. In the case of Subbarao, the images are made along exactly the same line, and so there are no angles that can be used to compute distance using triangulation.

It is also noted that there has been no assertion that Subbarao teaches conversion of images to make their pixel units equal, and that in fact Subbarao does not provide any teaching of that feature. Subbarao does teach certain types of preprocessing that are performed on images prior to using the images to determine distance (see col. 18, line 29 - col. 19, line 34). These preprocessing functions include sensor response correction, correcting for non-uniform light transmission, and brightness normalization. Changing the amount of object represented by each pixel is not taught. There appears to be no need to perform such conversion in Subbarao since Subbarao does not state or imply that his two cameras have different resolutions. As stated by Subbarao in reagard to the two-camera embodiment, the "image acquisition devices" (i.e. the photosensors) of each camera are the same as in the single camera embodiment, and therefore have the same resolution (col. 14, lines 22-25). Subarrao does not suggest that either the camera resolutions or the fields of view are different between the two cameras or between images made by the same camera. Therefore Subbarao does not and appears to have no motivation to perform pixel unit conversion on his images.

Obviousness analysis

It is well known that the obviousness analysis involved determining the differences between the cited references and the claimed invention, and then determining whether it would be obvious for one of ordinary skill to implement those differences in the manner claimed based on the teachings of the cited references.

The following table illustrates the differences between claim 7 (which is exemplary of all pending claims) and the teachings of Auty and Subbarao:

Claim 7	Auty	Subbarao
A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:		
a plurality of cameras for producing images of the object ***from different viewing angles***, the cameras having different resolutions from each other such that ***pixel units of the images are different in the amount of object represented thereby***;		1) Subbarao uses two cameras for producing images of an object from the same visual angle. 2) Subbarao's pixel units are not different in the amount of object represented thereby. Subbarao's images are made by the same camera, or by two different cameras that are not different with respect to their pixel units or fields of view.
a conversion unit for ***converting*** at least one of the images outputted from said plurality of cameras ***such that the pixel units of all images are equal in the amount of object represented thereby***; and	Auty converts the position of an object in the image plane to a position along the surface of a road in the image.	Subbarao's preprocessing does not involve converting images so that the pixel units of all images are equal, and nothing indicates that Subbarao's pixel units are unequal.
a depth image production section for processing the images ***using triangulation*** to calculate a distance to the object.	Auty determines the distance to an object using a single camera based on the position of the object within the image (i.e. its coordinates in the image plane).	Subbarao determines the distance to an object using two images of the object taken along the same line of sight.

From the table it is seen that the only similarity of either of the references to the present claims is that Auty uses two cameras having different resolutions. Both of the cited references fail to teach the following features that are required by every claim:

1) Converting images of an object taken from different viewing angles and having different resolutions so that their pixel units are the same in the amount of object represented.

2) Processing such images using triangulation to determine the distance to the object.

Given that both of the cited references lack any teaching of these two features, it is unreasonable to assert that the claimed invention would be obvious based on the teachings of the references. The references are superficially related to the claimed invention in that they involve determining distance, but otherwise they are completely different in their approaches. One following the teachings of the cited references would not arrive at the claimed invention since virtually all of the features of the claimed invention are both absent from and inapplicable to the technologies described in the cited references.

The foregoing remarks address all of the features of claim 7, and are equally applicable to claim 2 The remaining claims recite additional features in addition to those of claims 2 and 7 that are likewise not taught by the cited references. Claims 3 and 8 recite the additional features of using cameras having different fields and processing those images to have equal pixel units. Claims 5 and 9 recite the additional feature of converting images by selecting from among a set of samples of the image that are taken beginning at successive points of the image. Claims 6 and 10 recite the additional features of both claims 3 and 8 and claims 5 and 9. Claims 11, 13, 15 and 17 recite the feature that images are converted by sampling to reduce their resolution. Claims 12, 14, 16 and 18 recite the feature that images are converted by interpolation to increase their resolution. None of these features is taught or implied by the cited references.

Applicants therefore request the Board of Appeals to overturn the rejections of all claims based on the Auty and Subbarao references.

Respectfully submitted,

By Ronald Coslick

Ronald Coslick
Attorney for Applicant
Registration No. 36,489

Date: April 24, 2003

FOLEY & LARDNER
Washington Harbour
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5109
Telephone: (202) 672-5407
Facsimile: (202) 672-5399

APPENDIX

Claims on appeal:

2. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of cameras for producing images of the object from different viewing angles, the cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

conversion means for converting at least one of the images outputted from said plurality of cameras such that the pixel units of all images are equal in the amount of object represented thereby; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

3. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of first cameras for producing images of the object from different viewing angles, the first cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a plurality of second cameras for producing images of the object from different viewing angles, the second cameras having different visual fields from each other such that pixel units of the images are different in the amount of object represented thereby;

conversion means for converting at least one of the images outputted from said first cameras and at least one of the images outputted from said second cameras such that the pixel units of all images are equal in the amount of object represented thereby; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

5. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of cameras for producing images of the object from different viewing angles, the cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

conversion means for converting at least one of the images produced by said cameras such that the pixel units of all images are equal in the amount of object represented thereby, the conversion means converting the at least one image by selecting from among a set of samples of the at least one image that are sampled beginning at successive positions in the at least one image; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

6. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of first cameras for producing images of the object from different viewing angles, the first cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a plurality of second cameras for producing images of the object from different viewing angles, the second cameras having different visual fields from each other such that pixel units of the images are different in the amount of object represented thereby;

conversion means for converting at least one of the images outputted from said first cameras and at least one of the images outputted from said second cameras such that the pixel units of all images are equal in the amount of object represented thereby, the conversion means converting each of the at least one images by selecting from among sets of samples of each of the at least one images that are sampled beginning at successive positions in each of the at least one images; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

7. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of cameras for producing images of the object from different viewing angles, the cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a conversion unit for converting at least one of the images outputted from said plurality of cameras such that the pixel units of all images are equal in the amount of object represented thereby; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

8. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of first cameras for producing images of the object from different viewing angles, the first cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a plurality of second cameras for producing images of the object from different viewing angles, the second cameras having different visual fields from each other such that pixel units of the images are different in the amount of object represented thereby;

a conversion unit for converting at least one of the images outputted from said first cameras and at least one of the images outputted from said second cameras such that the pixel units of all images are equal in the amount of object represented thereby; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

9. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of cameras for producing images of the object from different viewing angles, the cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a conversion unit for converting at least one of the images produced by said cameras such that the pixel units of all images are equal in the amount of object represented thereby, the conversion means converting the at least one image by selecting from among a set of samples of the at least one image that are sampled beginning at successive positions in the at least one image; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

10. (Amended Four Times) A three-dimensional structure estimation apparatus which measures a distance to an object, comprising:

a plurality of first cameras for producing images of the object from different viewing angles, the first cameras having different resolutions from each other such that pixel units of the images are different in the amount of object represented thereby;

a plurality of second cameras for producing images of the object from different viewing angles, the second cameras having different visual fields from each other such that pixel units of the images are different in the amount of object represented thereby;

a conversion unit for converting at least one of the images outputted from said first cameras and at least one of the images outputted from said second cameras such that the pixel units of all images are equal in the amount of object represented thereby, the conversion means converting each of the at least one images by selecting from among sets of samples of each of the at least one images that are sampled beginning at successive positions in each of the at least one images; and

a depth image production section for processing the images using triangulation to calculate a distance to the object.

11. (New) The apparatus claimed in claim 2, wherein the conversion means samples images such that the pixel units of sampled images represent an amount of object represented by pixel units of an image having a lowest resolution.

12. (New) The apparatus claimed in claim 2, wherein the conversion means interpolates images such that the pixel units of interpolated images represent an amount of object represented by pixel units of an image having a highest resolution.

13. (New) The apparatus claimed in claim 3, wherein the conversion means samples images such that the pixel units of sampled images represent an amount of object represented by pixel units of an image having a lowest resolution.

14. (New) The apparatus claimed in claim 3, wherein the conversion means interpolates images such that the pixel units of interpolated images represent an amount of object represented by pixel units of an image having a highest resolution.

15. (New) The apparatus claimed in claim 7, wherein the conversion unit samples images such that the pixel units of sampled images represent an amount of object represented by pixel units of an image having a lowest resolution.

16. (New) The apparatus claimed in claim 7, wherein the conversion unit interpolates images such that the pixel units of interpolated images represent an amount of object represented by pixel units of an image having a highest resolution.

17. (New) The apparatus claimed in claim 8, wherein the conversion unit samples images such that the pixel units of sampled images represent an amount of object represented by pixel units of an image having a lowest resolution.

18. (New) The apparatus claimed in claim 8, wherein the conversion unit interpolates images such that the pixel units of interpolated images represent an amount of object represented by pixel units of an image having a highest resolution.